Forensic Analysis of Mitochondrial DNA: Application of Multiplex Solid-Phase - Fluorescent Minisequencing to High Throughput Analysis

G. Tully[1], J. M. Morley[1] and J. E. Bark[2]
[1]The Forensic Science Service, 109 Lambeth Road, London, SE1 7LP
[2]The Forensic Science Service, Priory House, Gooch Street North, Birmingham, B5 6QQ.

INTRODUCTION

Mitochondrial DNA (mtDNA) has been used as a tool for forensic identification since 1993 (1,2). It has now become widely accepted, its US courtroom debut being reported in both the scientific and popular press (3-5). Perhaps the greatest utility of mtDNA analysis in a forensic context is in the field of hair comparison. MtDNA can give valuable objective evidence where only subjective judgements were previously possible.

Despite the clear potential of mtDNA to enable elimination of hairs as originating from a single source, it cannot be employed in all cases involving hair comparison due to the time-consuming and therefore costly nature of sequencing techniques. Although robotic automation of sequencing reactions provides a solution to some of the labour-intensive stages, robotic workstations are cost-effective only for high throughput laboratories, and do not eliminate the analysis stages. Sequence analysis software is invaluable in the primary stages of analysis, but each base of each sequence must be manually inspected by an experienced operator in order to ensure that any possible heteroplasmy is detected.

The technique of Multiplex Solid-Phase Fluorescent Minisequencing (6) was developed to enable screening of the positions most likely to differ between individuals. Such a strategy minimises the number of samples for which sequencing is required, making the analysis in each case more efficient and enabling mtDNA analysis to be employed in a wider range of cases. The multiplex minisequencing technique has been used in forensic casework in the UK since January 1998, and this paper describes its use, and details case examples.

A complicating factor in forensic mtDNA analysis is the co-existence of more than one species of mtDNA molecule within an individual. This phenomenon, known as heteroplasmy, was originally thought to occur only in disease states, but was observed in the major non-coding region of mtDNA extracted from the putative remains of Tsar Nicholas II (7). It has subsequently been observed at many positions within the non-coding region (8-12) and differential segregation of the mtDNA species into different body tissues and indeed into different hairs within an individual has been documented (11,12). The implications of heteroplasmy and somatic mosaicism of mtDNA in forensic analysis are discussed and guidelines for reporting mtDNA evidence presented.

MATERIALS AND METHODS

Amplification

One to ten microlitres of DNA were amplified in a total volume of 50µl, using amplification conditions as published (6) with the modification that primer H00326 was not biotinylated. After amplification, 10µl of each PCR product was examined by agarose gel electrophoresis for the presence of correctly sized products.

PCR Product Immobilisation

Dynabeads M-280 streptavidin (Dynal, Oslo) were prewashed according to the manufacturer's instructions and resuspended in 40µl 2x binding and washing (BW) buffer (2M NaCl, 1mM EDTA and 10mM Tris-HCL, pH7.5). The remaining 40µl of each PCR product was bound to 40µl of Dynabeads by a 15 min incubation at 48^{o}C. The Dynabead-PCR product complex was washed in 2xBW and then in sterile distilled water (SDW). Following a 4 min incubation in 0.15M NaOH at room temperature to denature the PCR product and a wash in 0.15M NaOH, the samples were washed once in 250mM Tris-HCl, pH8, 0.1% Tween-20, once in 10mMTris-HCl, 1mM EDTA and once in SDW. All washes were carried out in 100µl volumes.

Minisequencing Extension Reactions

The bead-product complex was resuspended in 10 µl SDW. For each sample, 40µl of minisequencing extension multimix consisting of 0.4M Tris-HCl, 10mM $MgCl_2$ pH 9, 20% dimethyl sulphoxide, 0.1µl fluorescently labelled ddCTP and ddGTP, 0.125µl fluorescently labelled ddATP and ddTTP, 2.5 U AmpliTaq FS (Perkin-Elmer Limited, Warrington England) and the 13 primers listed in Table 1 was added to a fresh set of thin-walled tubes. The GeneAmp PCR System 9600 was pre-

heated to 57°C. The programme was paused and the tubes containing the multimix and those containing the bead-product complex were placed in the wells. The programme was resumed and samples and multimix were heated at 57°C for 30s. The programme was paused, the multimix added to the bead-product complex and the temperature reduced to 52°C for a 60s incubation. The bead-product complex was snap-cooled on ice and the multimix decanted. Samples were resuspended in 100µl 1xBW buffer.

Electrophoresis and Analysis

The 1xBW buffer was decanted and the samples resuspended in 4µl formamide and denatured at 72°C for 4 min before being snap-cooled on ice. The formamide, containing the labelled DNA was transferred to a clean tube and 1µl loaded on a 19% denaturing acrylamide gel with a well-to-detection distance of 12 cm in a 377 automated sequencer (Applied Biosystems Division). Electrophoresis was carried out on the 2xA setting (1680V) for 2 h and the data collected using ABI Prism 377xl Collection software (Applied Biosystems Division). The data were analysed using ABI Prism GeneScan Anaysis version 2.1 software (Applied Biosystems Division) and the peak colour for each extended minisequencing primer were scored for each sample.

RESULTS AND DISCUSSION

Casework Applications of Multiplex Solid-phase Fluorescent Minisequencing

Since its launch in January 1998, multiplex minisequencing of mitochondrial DNA has been used in cases ranging from murder and rape to robbery and drugs offences. To date, 16 cases have resulted in exclusion of at least one individual of interest as a possible source of a questioned sample, 7 have resulted in matches between a questioned sample and a suspect, and 8 have resulted in matches between a questioned sample and a victim. In addition, hairs from one putative series of rapes have been excluded as originating from a single individual, whilst hairs from a second putative series of rapes have been found to match each other, providing support for a link between the crimes. The success rate for analysis of shed hairs has been approximately 80%, and for faecal material, in excess of 90%.

The utility of mtDNA minisequencing is illustrated in the following case example. A hair was recovered from the underwear of a rape victim. This hair was microscopically similar to a reference sample from the suspect, and so was submitted for mtDNA analysis. Minisequencing was performed on the questioned hair, and on reference samples from the victim and the suspect. The results are illustrated in figure 1. The questioned hair differed from the suspect's sample only at position 73. This position is known to be stable, with no mutations having been observed in intergenerational (10) or somatic (13) mutation studies; it is also thought by phylogenetic analyses to be an ancient polymorphism, which has not mutated frequently (14). The suspect was thus excluded as the source of this hair. The profile from the questioned hair matched that of the victim; this profile had been observed 10 times in a database of 152 British Caucasians, so the match was assessed as providing moderate support for the proposition that the questioned hair originated from the victim.

Minisequencing of mtDNA thus provided valuable exclusionary evidence which had not been provided by microscopic analysis in this case. Sequencing did not prove necessary, thus enabling the case to be analysed more quickly and efficiently, at a lower cost.

Reporting Guidelines for Mitochondrial DNA Evidence

It is accepted that somatic (11,12) and germline (eg. 10) substitutions occur in mitochondrial DNA. Interpretation of mitochondrial DNA evidence must therefore incorporate this knowledge in order to minimise the chance of falsely excluding an individual as the source of an unknown sample.

As a starting point, we know that not all types of mutation are equally likely. Transitions are the most common mutation type, length mutations the second most common type and transversions the least common type of mutation (eg 15). It is also known that long homopolymer tracts are particularly prone to length mutations and heteroplasmy (16). Furthermore, not all tissues are equally likely to show 'fixed' somatic substitutions (12,13), where a 'fixed' substitution is defined as a change in the predominant haplotype from one state to another such that no heteroplasmy is observed using direct sequencing or minisequencing methodologies. Studies in our laboratory have revealed no differences between blood, saliva, faeces, and semen or vaginal material within individuals, but fixation of substitutions in individual hair shafts has been observed. It is known that not all positions within the mitochondrial DNA are equally prone to substitutions (eg. 17); positions 16093 and the length polymorphism at 309 appear to be substitution hotspots, with multiple observations of substitutions (eg. 10,11,18,19).

In our laboratory, more than one difference between samples from an individual has not been observed. Therefore, as a general guideline, a single base difference

between an unknown and a reference sample is not used to conclusively eliminate the donor of the reference sample as the source of the unknown sample. Again in general, two differences between a questioned and a known sample may be said to provide limited or moderate support for the samples having different origins, and three differences would represent moderate or strong support for the samples having different origins. However, the background knowledge of the occurrence of mutations of different types at different positions within the mtDNA genome, as described above, must be superimposed onto the general guidelines. In this way, all available information is used to estimate the likelihood of two samples originating from different individuals. Germline mutation occurs once in approximately 30 generations (10), and the number of generations separating the donor of a reference sample from the putative donor of an unknown sample must therefore be considered.

The requirement for multiple differences between samples in order to confirm an elimination would greatly reduce the efficacy of a screening method such as multiplex minisequencing. However, the rapid nature of this technique has enabled more data to be accrued on the occurrence of somatic mutations at the twelve positions examined using this test (13,20), and more stringent guidelines can therefore be formulated. In general, if the only difference between two samples is at a position of length variation, sequencing would be performed in order to clarify whether or not the samples may have originated from a single individual. Furthermore, the average number of pairwise differences between Caucasians using this technique is >3, calculated from a database of 152 British Caucasians. In the majority of cases, therefore, samples which have different maternal origins would be distinguished by more than one base.

The occurrence of heteroplasmy in both a known and an unknown sample at the same position would add weight to the strength of evidence associating the two samples. As more data on the occurrence of heteroplasmy are accrued, it may be possible to provide a numerical estimate of this increase in evidential weight.

In summary, each case must be treated on its own merits, using the guidelines suggested above to aid interpretation. In our experience, exclusions are often extremely clear-cut with 6 or more differences between samples. Often, more than one unknown sample is examined, and sequencing multiple hairs from an individual is an option should any doubt remain. Mitochondrial DNA analysis is thus a reliable and invaluable tool for forensic examination, enabling objective assessment of the likelihood of two samples originating from individuals on the same maternal lineage.

REFERENCES

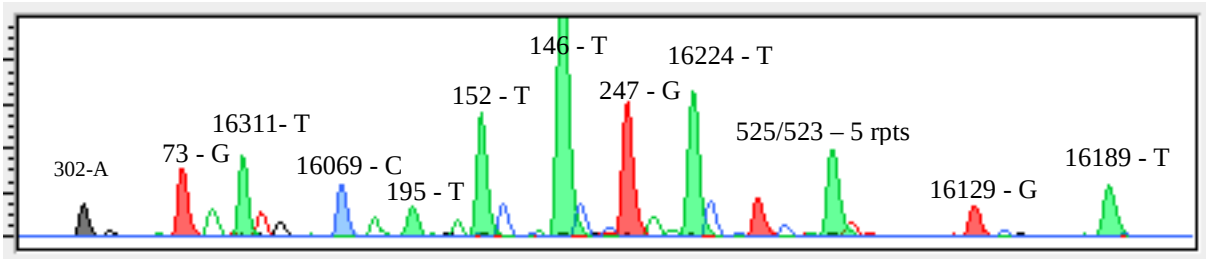
1. Holland M.M., Fisher D.L., Mitchell L.G., Rodriquez C., Canik J.J., Merril C.R., Weedn V.W.. Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam war. *J Forensic Sci 1993;* **38**:542 – 553.
2. Wilson M.R., Stoneking M., Holland M.M., DiZinno J.A.. Guidelines for the use of mitochondrial DNA sequencing in forensic science. *Crime Lab Digest 1993*; **20**:68 – 77.
3. Marchi E., Pasacreta R.J.. Capillary electrophoresis in court: the landmark decision of *The People of Tennessee versus Ware*. *J Cap Elec 1997*; **4**:145-156.
4. *The People of Tennessee versus Paul Ware*, 11 Judicial Dist. 203757, Div. 1 Criminal Ct. (1996).
5. The Chattanooga Times vol CXXVII, no 226, Sept 4, 1996:1. Ware found guilty; could get death.
6. Tully G., Sullivan K.M., Nixon P., Stones R.E., Gill P. Rapid detection of mitochondrial sequence polymorphisms using multiplex solid-phase fluorescent minisequencing. *Genomics 1996*; **34**:107 – 113.
7. Gill P., Ivanov P.L., Kimpton C., Piercy R., Benson N., Tully G., Evett I., Hagelberg E., Sullivan K. Identification of the remains of the Romanov family by DNA analysis. *Nat Genet 1994*; **6**:130 – 135.
8. Comas D., Paabo S., Bertranpetit J. Heteroplasmy in the control region of human mitochondrial DNA. *Genome Research 1995*; **5**:89 – 90.
9. Bendall K.E., Macaulay V.A., Baker J.R., Sykes B.C.. Heteroplasmic point mutations in the human mtDNA control region. *Am J Hum Genet 1996*; **59**:1276 – 1287.
10. Parsons T.J., Muniec D.S., Sullivan K., Woodyatt N., Alliston-Greiner R., Wilson M.R., Berry D.L., Holland K.A., Weedn V.W., Gill P, Holland M.M.. A high observed substitution rate in the human mitochondrial DNA control region. *Nat Genet 1997*; **15**:363–368.
11. Sullivan K.M., Alliston-Greiner R., Archampong FIA, Piercy R., Tully G., Gill P., Lloyd-Davis C. A single difference in mtDNA control region sequence observed between hair shaft and reference samples from a single donor. In: Proceedings from the Seventh International Symposium on Human Identification; 1996 Sept 19-21; Scottsdale (AZ):1996; Promega Corporation, pp 126–129.
12. Wilson M.R., Polanskey D., Replogle J., DiZinno J.A., Budowle B. A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes. *Hum Genet 1997*; **100**:167–171.
13. Morley J.M., Bark, J.E., Evans C.E., Perry J.G., Hewitt C.A., Tully G. Validation of mitochondrial DNA minisequencing for forensic casework. Submitted to Int J Leg Med.
14. Wilkinson-Herbots H.M., Richards M.B., Forster P., Sykes B.C. Site 73 in hypervariable region II of the human mitochondrial genome and the origin of European populations. *Ann Hum Genet 1996*; **60**:499-508.
15. Horai S., Hayasaka K.. Intraspecific nucleotide sequence differences in the major non-coding region of human mitochondrial DNA. *Am J Hum Genet 1990*; **46**:828-842.
16. Bendall K.E., Sykes B.C. Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *Am J Hum Genet 1995*; **57**:248-256.
17. Wakeley J. Substitution rate variation among sites in hypervariable region I of human mitochondrial DNA. *J Mol Evol 1993*; **33**:613-623.
18. Parsons T.J. and Holland M.M., personal communication.
19. Alliston-Greiner R *et al*., manuscript in preparation.
20. Tully G., Robilliard D., Barry D., unpublished results.

ACKNOWLEDGEMENTS

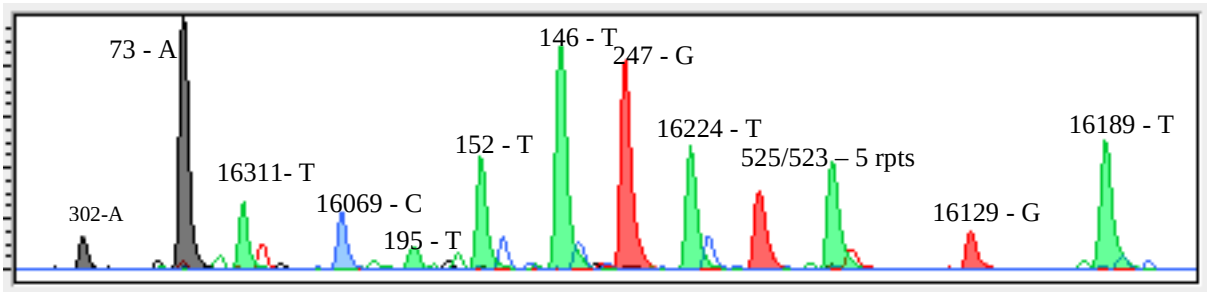
The input of all members of the mitochondrial DNA research and casework teams is gratefully acknowledged.

Location of 3' base	Sequence 5' - 3'	Concentration (µM)
H00303	GTC GTG TGG CCA GAA GCG GGG GGA GGG GGG G	6
H0074	TTT TTT TTT CCA GCG TCT CGC AAT GCT ATC GCG TGC A	0.3
H16312	TTT TTT TTT TGT GCT ATG TAC GGT AAA TGG CTT TAT GT	0.06
H16070	TTT TTT TTT TTT TTT TAA GCA TTA ATT AAT TAA CAC RCT TYR RTA	0.4
H00196	TTT TTT TTT TTT TTT TTA AGC ATT AAT TAA TTA ACA CRC TTY RRT A	6.5
H00153	TTT TTT TTT TTT TTT TCT GTA ATA TTG AAC GTA GGT GCG ATA AAT AAT	0.3
H00147	TTT TTT TTT TTT TTT TTA ATA TTG AAC GTA GGT GCG ATA AAT AAT RRR ATG	1.25
H00248	TTT TTT TTT TTT TTT TTT TTT TTT TTC TGT GTG GAA AGT GGC TGT GCA GAC ATT	0.05
H16224	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTG GAG TTG CAG TTG ATG TGT GAT AGT TG	0.1
L00522	TTT TTT TTT TTT TTT TTT TCT CAT CAA TAC AAC CCC CGC CCA TCC TAC CCA GCA CAC ACA CAC	0.6
LOO524	TTT TTT TTT TTT TTT TTT TTT TTT CTC ATC AAT ACA ACC CCC GCC CAT CCT ACC CAG CAC ACA CAC	0.48
H16130	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TGT ACT ACA GGT GGT CAA GT	0.04
H16190	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT GGT TGA TTG CTG TAC TTG CTT GTA AGC ATG RGG	0.2

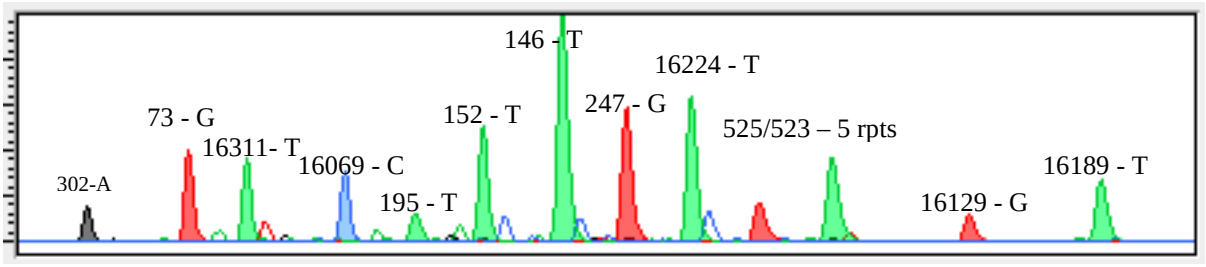
Table 1 Minisequencing Extension Primer Sequences



Hair recovered from underwear of victim



Reference sample from suspect



Reference sample from victim

Figure 2: Case example using multiplex minisequencing; the colour of each peak represents the base which has been incorporated. The hair recovered from the underwear of the victim differs from the reference sample from the suspect at position 73.